

Modelagem computacional no diagnóstico de Transtorno do Espectro Autista.

Palavra-chave: autismo, diagnóstico, simulação, modelagem

Resumo

O Seguinte estudo visa o desenvolvimento de um modelo de simulação computacional para atuar no diagnóstico de Transtorno do Espectro Autista (TEA). Sendo assim inicialmente escolheu-se um modelo que pudesse ser aplicado na problemática, e dessa forma foram realizados testes com intuito de avaliar o desempenho do programa bem como sua possível aplicação em um cenário real.

Introdução/justificativa

Por definição o Transtorno do Espectro Autista (TEA) configura-se como um distúrbio do neurodesenvolvimento no qual o indivíduo apresenta manifestações comportamentais atípicas, tais como déficits na comunicação e interação social, padrões de comportamentos repetitivos e estereotipados, atraso na fala durante os primeiros anos de vida, entre outros aspectos.

Vale salientar que o autismo configura-se como um termo recente na literatura médica, pois apenas em meados de 1910 o psiquiatra e cientista Eugen Bleuler aderiu o termo “Autismo” para classificar pessoas que apresentavam dificuldade de interação social e isolamento. Desse modo, durante muito tempo o autismo foi visto como um tabu dentro o âmbito social, já que a falta de um diagnóstico adequado, aliado ao baixo entendimento acerca do assunto, contribuiu para que muitas pessoas com quadro de TEA fossem negligenciadas ou recebessem o tratamento inadequado ao longo dos anos.

Nesse sentido, outro fator que deve ser levado em consideração é a importância do diagnóstico apropriado, tendo em vista que o autismo pode ser classificado em subcategorias as quais variam de acordo com o nível de intensidade dos sintomas de TEA no indivíduo. Logo torna-se difícil tanto para a pessoa com espectro autista quanto para seus familiares, amigos ou pessoas próximas identificarem a existência do distúrbio, dificultando portanto o processo de buscar auxílio e acompanhamento médico adequado.

Objetivo

O objetivo do seguinte projeto é auxiliar na identificação e diagnóstico do espectro autista.

Objetivos específicos

- (1) - Elaborar um modelo computacional que possa avaliar uma base de dados, a qual fora desenvolvida por meio de uma pesquisa do espectro autista em indivíduos de diferentes faixas etárias, e a partir dessa análise gerar a simulação de um cenário de identificação do espectro autista;
- (2) - Aplicar o sistema de Redes neurais artificiais ao modelo;
- (3) - Avaliar a acurácia dos resultados obtidos mediante a geração de gráficos;
- (4) - Classificar a confiabilidade do modelo baseado em seus resultados;

Metodologia

Tendo em vista o objetivo de pesquisa deste projeto a metodologia usada consiste no desenvolvimento de um modelo de simulação computacional baseado em redes neurais artificiais (MLP). Sendo assim, para a execução dos testes utilizou-se a base de dados desenvolvida pela Olympiad - Autism Prediction Challenge (Kaggle Community Competitions, associado e hospedado pela ML GDEs, e patrocinado por Google Developers). O conjunto de dados base deriva de uma pesquisa realizada para mais de 700 pessoas que preencheram um formulário de aplicativo. Na elaboração dos testes foram selecionadas as características relevantes da base de dados de acordo com o modelo de simulação proposto.

Sendo eles os seguintes:

- ID- Identificação do paciente
- A1_Score to A10_Score- Pontuação baseada na ferramenta de triagem de 10 itens do Autism Spectrum Quotient (AQ)
- gender- Sexo do paciente
- jaundice- Se o paciente teve icterícia no momento do nascimento
- autism- Se um membro imediato da família foi diagnosticado com autismo
- result- Pontuação para o teste de triagem AQ1-10
- Class/ASD- Resultado classificado como 0 ou 1. Aqui 0 representa Não e 1 representa Sim.

Além disso, para a avaliação do desempenho da rede, os testes foram realizados separando a base de dados inicialmente com 70% para treino e 30% para teste, após isso 80% para treino e 20% para teste e por fim 90% para treino e 10% para teste. O número de neurônios também foi alterado ao longo dos testes, para fins de melhor desempenho, e a função de ativação utilizada foi a “relu”, a qual fora escolhida devido sua praticidade de processamento.

Resultados esperados

Espera-se que a partir do treinamento da rede o modelo consiga gerar resultados nas amostragens de teste para que dessa forma seja possível avaliar os resultados baseado na acurácia do modelo, ou seja, se o modelo conseguiu realizar a classificação apropriada dos dados submetidos a amostragem. Em caso de bom desempenho espera-se que seja possível avaliar a possibilidade de utilizar a rede em cenários que possam contribuir para a resolução da problemática.

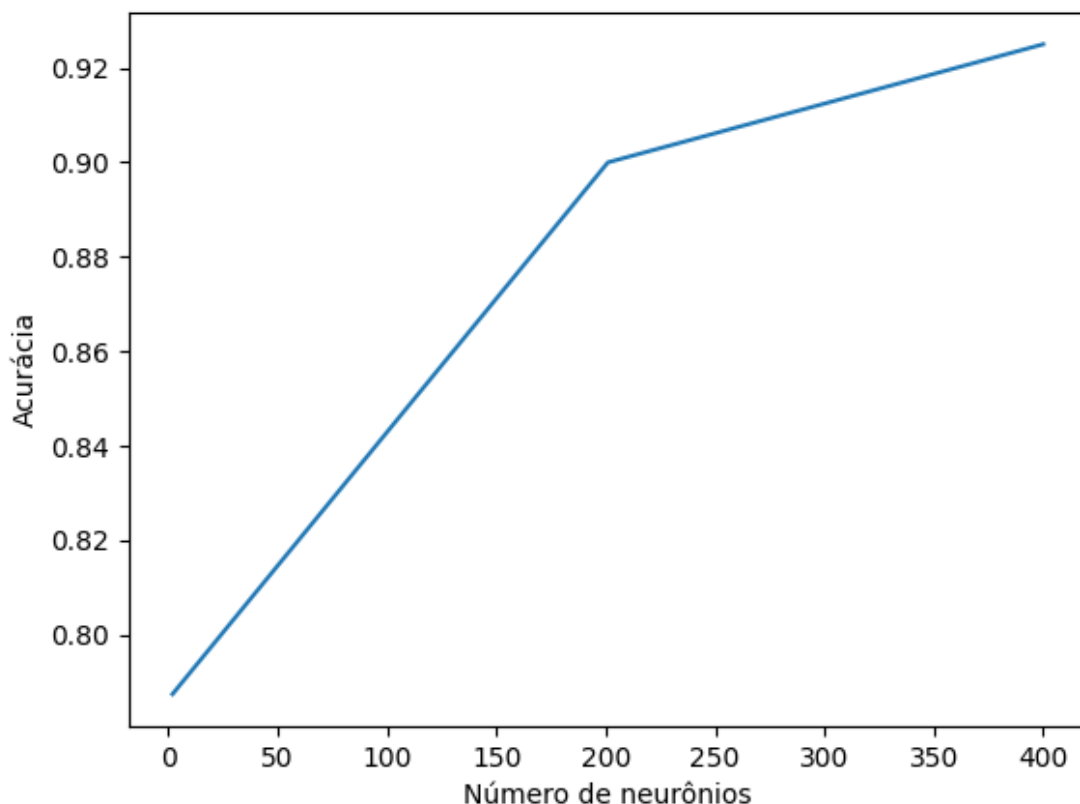
Resultado

Mediante os casos de treino e teste o qual o modelo fora submetido, foram gerados as matrizes de confusão para cada amostragem, a fim de avaliar a acurácia do modelo no conjunto de dados, dessa forma obtiveram-se os seguintes resultados:

```
Acurácia 70/30: 0.833
Matriz de confusão para 70/30:
[[168  12]
 [ 28  32]]
Acurácia 80/20: 0.844
Matriz de confusão para 80/20:
[[105  17]
 [  8  30]]
Acurácia 90/10: 0.925
Matriz de confusão para 90/10:
[[62  1]
 [ 5  12]]
```

Logo, observou-se que a medida que a amostragem de treino foi aumentada, (mudando de 70% para 80% e por fim 90%) a acurácia do modelo sofreu elevações positivas. Dessa forma, podemos concluir que o desempenho da rede melhora à medida que reforçamos seu treinamento.

Ademais, foi gerado um gráfico para avaliar o desempenho do modelo mediante a mudança no número de neurônios:



Observa-se que a adição no número de neurônios na rede gerou um efeito benéfico para a acurácia obtida, esse cenário pode ser derivado pelo aumento de sinapses a medida que

adiciona-se neurônios, fazendo com que a taxa de aprendizado do modelo melhore. Portanto, podemos concluir que a acurácia do modelo possui uma relação positiva com o aumento no número de neurônios da rede.

Discussão

Mediante os resultados obtidos pelo modelo podemos concluir que sua eficácia no processo de simulação depende da combinação adequada de fatores que contribuam para a taxa de aprendizagem da rede, como o acréscimo do número de neurônios e o tamanho da amostragem para treinamento, vale ressaltar que outros parâmetros, como mudança na função de ativação e número de interações dos neurônios, também precisam ser levados em consideração para obter-se um melhor desempenho do modelo.

Portanto, dado o objetivo principal do projeto, que visa a simulação no diagnóstico do espectro autista, os resultados obtidos contribuem de forma benéfica para a causa, todavia o aperfeiçoamento do modelo faz-se necessário, para que assim este possa ser refinado de modo a buscar resultados mais precisos.

Referências:

- <https://www.kaggle.com/competitions/autism-prediction/data>
- <https://github.com/MichelleTxr/simulacao-tea.git>
- <https://www.saude.pr.gov.br/Pagina/Transtorno-do-Espectro-Autismo-TEA>
- <https://drauziovarella.uol.com.br/doencas-e-sintomas/transtorno-do-espectro-autista-tea/>